May 2002

Calculation of Charge-Based Relative Values for Laboratory Procedures

Final Report

CMS Contract No. 500-95-0061/T.O. #2

Prepared for

Michael Park
Centers for Medicare & Medicaid Services
Mail Stop C4-17-27
7500 Security Boulevard
Baltimore, MD 21244-1850

Prepared by

University of Wisconsin–Madison
Center for Health Systems Research and Analysis
610 Walnut Street
Madison, WI 53705

RTI
Division of Health Economics Research
3040 Cornwallis Road
Research Triangle Park, NC 27709

Northwestern University
629 Noyes Street
Evanston, IL 60208

CHSRA CENTER FOR HEALTH SYSTEMS RESEARCH & ANALYSIS

RTI INTERNATIONAL

Calculation of Charge-Based Relative Values for Laboratory Procedures

Final Report

CMS Contract No. 500-95-0061/T.O. #2

May 2002

Prepared for

Michael Park
Centers for Medicare & Medicaid Services
Mail Stop C4-17-27
7500 Security Boulevard
Baltimore, MD 21244-1850

Prepared by

Sarita L. Karon, Ph.D.
University of Wisconsin–Madison
Center for Health Systems Research and Analysis
610 Walnut Street
Madison, WI 53705

Thomas J. Hoerger, Ph.D., Project Director
Lee Rivers Mobley, Ph.D.
RTI
Division of Health Economics Research
3040 Cornwallis Road
Research Triangle Park, NC 27709

Richard C. Lindrooth, Ph.D.
Northwestern University
629 Noyes Street
Evanston, IL 60208

Contents

Figures

Tables

EXECUTIVE SUMMARY

ES.1 Background and Purpose of the Report

The Institute of Medicine (IOM) recently undertook an exhaustive assessment of the current payment system used for reimbursement of outpatient laboratory procedures by Medicare Part B. In its report, *Medicare Laboratory Payment Policy: Now and in the Future* (2000), the IOM concluded that the current fee schedule, which is based in part on laboratory charges from 1983, provides sufficient beneficiary access but does not provide appropriate and flexible mechanisms for making changes in fees for individual procedures. The IOM recommends that Medicare payments for outpatient clinical laboratory services should be based on a single rational national fee schedule. The building blocks for this system would be

- a relative value scale (RVS);
- a dollar conversion factor that transfers relative values into payment amounts;
- adjustments for laboratory, beneficiary, or other characteristics, including geographic location; and
- periodic updates.

The IOM report recommends that, on an interim basis, relative payment amounts should be based on the current national limitation amounts (NLAs), which constrain fees in most regions in the country. In the longer run, the report recommends that a data-driven consensus process for refining the fee schedule should be developed. The report recommends that the Centers for Medicare & Medicaid Services (CMS) should explore alternative methods for gathering data and identifies four approaches that merit further consideration:

- *Microcosting studies* to determine the costs of individual procedures. This approach could be used to set both the RVS and the conversion factor.
- *Competitive bidding demonstration* to set the RVS but not the conversion factor.
- *Negotiated fee demonstration* to set both the RVS and the conversion factor or just the RVS.
- *Analysis of charges* to set the RVS but not the conversion factor.

The purpose of this report is to analyze charge data, develop charge-based relative values, and provide comparisons of reimbursement across several hypothetical fee schedules. The comparisons show how fees would change if a new price schedule were to be based on charge-based relative values, and we explore how sensitive the relative values are to various partitions of the charge data. These analyses are meant to inform CMS about properties of the existing charge data, as a point of departure for possible future development of a consensus-based RVS. Development of the final RVS is beyond the scope of this report.

ES.2 Justification for a Charge-Based Approach

Ideally, the new RVS to be used in the Medicare laboratory fee schedule would accurately reflect the relative costs of different procedures. If procedure A costs twice as much to produce as procedure B, then procedure A's relative value should be twice as high as procedure B's relative value. The rationale for using charges to determine relative values is that charges may be systematically related to costs.

Under certain assumptions, economic theory suggests that laboratories will mark up prices over marginal costs by a percentage that depends on the elasticity of demand for a procedure. If the elasticity of demand is similar across procedures, the percentage markup will be the same for each procedure, and the ratio of prices between procedures will equal the ratio of relative costs. In this case, charge-based relative values will provide a useful measure of relative costs.

As mentioned, the preceding analysis relies on a number of assumptions, and the result that relative prices reflect relative costs may not hold if the assumptions are violated. These violations could occur in several ways. Elasticities may differ across procedures, leading to different markups. Marginal costs may be a function of the extant fee schedule, to the extent that price distortions in the fee schedule distort technological investment and diffusion. Laboratories may adopt more sophisticated pricing strategies than simple economic theories suggest or, at the other extreme, not set charges (list prices) in a systematic way because these charges are seldom used to set reimbursement rates.

Despite the potential for violations of the assumptions, we believe that analyzing charges is still a useful starting point in the development of a consensus-based RVS. The assumption that laboratories base charges in part on costs is reasonable because even if laboratories do not know the per unit costs for procedures, they most likely have a knowledge of the relative costs. Because of the potential violations, we recommend that charge data be used as one of the sources in the development of an RVS but not the only source. Charge data could be used as the starting point for a consensus panel of experts who can make adjustments based on their expertise or experience, or as a benchmark for comparing the results of other approaches for developing an RVS.

ES.3 Methodology

We analyze the most recent available (2000) Physician/Supplier Procedure Summary Master file. The file includes information from all Part B claims submitted to Medicare carriers for reimbursement. Because we want to show how reimbursements would shift if a new charge-based pricing system were to be implemented, we use as the comparison the current NLA-based pricing system. This is actually a blend of NLAs and regional fee schedule amounts because in the current system, reimbursement equals the lower of the NLA amount, the national fee schedule amount, or the submitted charge.

To transform the charge data into relative values, we divide each charge by a standard amount, or numeraire. In our main analysis, we use the weighted average charge for all procedures as the numeraire. The numeraire itself has a relative value of one (equal to the weighted average charge divided by the weighted average charge)—hence the term numeraire. The relative value of each procedure will then be expressed as a number that is either greater or less than one. We also perform subanalyses for 12 different procedure classes. In these subanalyses, the numeraire equals the weighted average charge for the procedures in the class.

We compare the charge-based relative values to relative charges based on the current NLA-based reimbursement system. Both graphical and statistical comparisons are presented. We also perform subanalyses to determine whether the calculated relative values vary across classes, between independent laboratories and physician office laboratories, across regions, or by volume.

After examining relative values, we calculate the prices that would hold if charge-based relative values were multiplied by a conversion factor that set total allowed charges equal to the actual allowed charges in 2000. The resulting "expenditure-neutral" fees allow us to compare the charge-based fees to current fees. Expenditure-neutral prices are calculated using both the universal numeraire and the set of class-specific numeraires. Using expenditure-neutrality and the universal numeraire, total allowed charges equal the current level, but allowed charges in each class may be less than or greater than current allowed charges. In contrast, using expenditure-neutrality and class-specific numeraires produces both total and class-specific neutrality.

ES.4 Results of the Analysis

Looking at all procedures with a universal numeraire, we find that the distribution of charge-based relative values is, from a statistical standpoint, significantly different from the distribution of relative values implied by the current NLA-based reimbursement system. There are a large number of outliers between the two sets of relative values, where outliers are defined as those procedures with unusually large percentage differences between the charge-based and NLA-based relative values. About 22 percent of procedures have a percentage difference that exceeds the standard deviation ($\pm$63 percent) from the mean difference (19 percent). This implies that prices for the outlier procedures would either decrease by at least 44 percent or increase by at least 82 percent if charge-based relative values were adopted in place of NLA-based relative values (assuming a constant conversion factor).

Applying class-specific numeraires to calculate relative values separately for each class does not eliminate the general difference in distributions between the charge-based and NLA-based relative values. Other subanalyses show that low volume procedures account for a disproportionate share of outliers, charge-based relative values calculated for independent laboratories are significantly different from the charge-based relative values calculated for physician office laboratories, and the distribution of charge-based relative values calculated for 5 of the 10 CMS regions differ from the charge-based relative values calculated for the nation as a whole.

ES.5 Comparison of Various Payment Systems

Replacing the current NLA-based reimbursement system with prices based on charge-based relative values would produce relatively small redistributions of payments across classes of procedures. Looking at individual procedures, prices based on charge-based relative values are frequently much different from the NLA-based prices, as our results for relative values would suggest. There are generally smaller differences in prices between the charge-based prices with universal and class-specific numeraires than there are between either type of charge-based prices and the NLA-based prices.

ES.6 Summary, Discussion, and Areas for Future Work

Calculating relative values for nearly 1,000 clinical laboratory procedures necessarily produces an abundance of relative values for individual procedures, and amidst this abundance it is easy to lose track of the big picture. To focus our summary, we concentrate on answering five general questions.

Would charge-based relative values be significantly different from the relative values implied by the current NLAs?

Yes. Based on our statistical tests, the distribution of charge-based relative values is significantly different from the distribution implied by the NLAs. Moreover, looking at individual procedures, there are a number of large “outliers,” which represent substantial price differences.

Does it matter if you use a universal numeraire instead of a class-specific numeraire to compute the relative values?

Probably not. Our results suggest that the choice of numeraire does not have a large effect on the relative values we compute. For most procedures, the relative values computed with universal and class-specific numeraires are very close; they are almost always closer to each other than either of them is to the relative value associated with NLAs. The list of procedures that are outliers, relative to NLA-based relative values, is very similar for the universal and class-specific numeraires. Finally, the calculation of expenditures suggests that there will be relatively little redistribution between classes if a universal numeraire is adopted.

Would an expenditure-neutral pricing system based on charge-based relative values change the distribution of reimbursements, relative to the current NLA-based system?

At more aggregate levels, payments under an expenditure-neutral pricing system based on charge-based relative values would lead to relatively little redistribution across procedure classes, between independent and physician office laboratories, and across regions, as measured relative to the current NLA-based system. However, at the individual procedure level, changing from NLA-based reimbursement to charge-based reimbursement would lead to large changes in reimbursement for individual procedures. Relatively few procedures would have small differences (where the charge-based expenditure is within 10 percent of the NLA-based expenditure). At the extremes, charge-based reimbursement would be more than 33 percent less than NLA-based reimbursement for 15 percent of procedures, and greater than 33 percent more than NLA-based reimbursement for 28 percent of procedures. However, these procedures account for a much smaller share of current volume and allowed charges. Thirty-three procedures would have reduced reimbursement of at least $1 million under a charge-based system, while 26 procedures would each receive increased reimbursement of at least $1 million.

If charge-based relative values were adopted, would utilization of individual tests change?

Our calculation of expenditure-neutral prices implicitly assumes that the volume of utilization of individual tests will not change if we move from NLA-based reimbursement to reimbursement on the basis of charge-based relative values. In practice, given the large changes in prices for individual procedures associated with the charge-based relative values, it is possible that volume will change for some procedures. Laboratories might increase marketing for tests with higher prices and decrease marketing for tests with lower prices. The net effect of volume changes on overall laboratory expenditures is not clear.

Did our analysis identify any other problems associated with charge-based relative values?

We found that calculating charge-based relative values may not help set reimbursement rates for automated test panels (ATPs). In the current fee schedule, selected automated chemistry tests are bundled for payment purposes, with reimbursement based purely on the number of tests performed and payments rising less than proportionately with the number of tests. The theoretical assumptions used to link relative prices and relative costs are more likely to be violated for these tests. That could be a problem for applying charge-based relative values, because panel tests and individual test components of ATPs account for nearly 40 percent of Medicare allowed charges.

Future Work

Our results indicate a number of areas for future work if CMS decides to proceed with incorporating charge data into an RVS. These areas include validation, applicability to hospital outpatient laboratories, transition issues, panel tests, and new tests.

1. INTRODUCTION

Medicare is the largest payer for clinical laboratory procedures in the nation, covering inpatient and outpatient testing for the elderly and disabled. The Institute of Medicine (IOM) recently undertook an exhaustive assessment of the current payment system used for reimbursement of outpatient laboratory procedures by Medicare Part B. Medicare Part B covers outpatient laboratory procedures performed by independent laboratories, physician office laboratories, hospital outpatient departments, and other noninpatient facilities. Current payment rates for outpatient procedures are set according to a prospective system based on 1983 customary charge data and implemented in 1984 with laboratories being paid the lesser of submitted charges or a fee schedule amount. Initially, payment rates under the fee schedules were set separately in each of 56 geographic jurisdictions, limited by a national cap. The 75th percentile of 1983 prevailing charges defined the fee schedule amounts in each of the 56 areas, and a mechanism was used to update the fees annually, based on the change in the Consumer Price Index (CPI). For most years, however, Congress has specified lower update factors. Currently, national caps called National Limitation Amounts (NLAs) set ceilings on payment rates for most procedures. Although the NLAs constrain most fees in most areas, the 56 separate fee schedules are still operational. Thus, current reimbursement is the lesser of the submitted charge, regional fee schedule, and the NLA payment rate. Throughout the report, we refer to this current reimbursement amount as the "current NLA-based" system.

Constraints on payments have contributed to a decline in actual Medicare expenditures for laboratory procedures, while expenditures for most other medical services have continued to rise. The laboratory payment system has evolved gradually over the past two decades with key decisions regarding coverage, payment, and medical necessity made both nationally and locally by private Medicare administrators. There has not been a formal methodology guiding these decisions, and it is unclear whether the resulting fees accurately reflect the costs of providing individual tests. Concerns about how well Medicare reimbursements reflect current costs of laboratory testing and about the system's ability to keep up with anticipated changes in technology prompted Congress to direct the Centers for Medicare & Medicaid Services (CMS) to commission the IOM study. The IOM report *Medicare Laboratory Payment Policy: Now and in the Future* (2000) concludes that the current Medicare payment policy provides adequate beneficiary access but does not provide more appropriate and flexible mechanisms for making changes in payments for individual tests. The report recommends that Medicare payments for outpatient clinical laboratory services should be based on a single rational national fee schedule. The building blocks for this system would be

- a relative value scale (RVS);
- a dollar conversion factor that transfers relative values into payment amounts;
- adjustments for laboratory, beneficiary, or other characteristics, including geographic location; and
- periodic updates.

The report recommends that, on an interim basis, relative payment amounts should be based on the current NLAs. In the longer run, the report recommends that a data-driven consensus process for refining the fee schedule should be developed. The report recommends that CMS should explore alternative methods for gathering data and identifies four approaches that merit further consideration:

- *Microcosting*. Studies would be conducted to determine the costs of individual procedures. This approach could be used to set both the RVS and the conversion factor, or just the RVS.
- *Competitive bidding demonstration*. In principal, this approach could be used to set both the RVS and the conversion factor. However, the IOM report recommends using this approach only to set the RVS.
- *Negotiated fee demonstration*. This approach could be used to set both the RVS and the conversion factor or just the RVS.
- *Analysis of charges*. Under this approach, charge data would be used to set the RVS but not the conversion factor.

In response to a Task Order issued by CMS in 1998, the University of Wisconsin–Madison, RTI, and Northwestern University were awarded a contract to design and implement a competitive bidding demonstration for Medicare laboratory procedures. After the publication of the IOM study, our contract was modified so that we were then charged with developing charge-based relative values for Medicare laboratory procedures.

In this analysis, we calculate and analyze charge-based relative values using data from the most recent available (2000) Physician/Supplier Procedure Summary File. In Section 2, we discuss the advantages and disadvantages of using charge-based data as a point of departure for possible future development of a final, consensus-based RVS. We stress that development of a final RVS is beyond the scope of this analysis. In Section 3, we first describe some measures taken to ensure reliability/reduce noise in the raw submitted charges data. We then describe the methodology used in calculating the charge-based relative values and the approach used in comparing the charge-based relative values to those implied by the existing NLA-based system. The NLA-based relative values reported here are actually a blend (the lesser) of the NLA fee schedule, submitted charges, and regional fee schedules.[1] We use in this report the phrase "NLA-based" to mean the current system that represents actual payments.

In the main analysis, we use the Kolmogorov–Smirnov test (described in Appendix A) to compare the two relative value distributions and test the hypothesis that the two distributions are the same. We provide tables with this test statistic and other descriptive information, such as the proportion of procedures whose differences across the two schedules were more than one standard deviation from the mean.

A graph is used to plot the two relative values for each procedure (by HCPCS code) against one another. With a perfect match, the plot would be a straight line emanating from the origin with a slope of 1; the degree of dispersion in the scatter from the hypothetical straight line is a visual representation of the discrepancy between the two relative value schedules across all procedures.

In Section 4, we discuss the findings from this main analysis, then turn to sensitivity testing to seek additional information about the observed discrepancies. We report the findings from the sensitivity tests, which include subanalyses performed to calculate

[1]In current data, submitted charges are very rarely lower than the other two fee schedules.

separate charge-based relative values for different subsets of the data. These subsets are procedure volume, procedure class, laboratory type, and region. We provide descriptive tables, Kolmogorov–Smirnov tests (where applicable), and graphic plots for these subanalyses, similar to those provided for the main analysis.

We use the results of the sensitivity tests to determine whether relative values based on subsets of the data are significantly different from those calculated using all of the data. In the next part of the analysis, we construct several hypothetical pricing schemes for comparative purposes (Section 5). We describe properties of an expenditure-neutral pricing system derived from the charge-based relative values and discuss how current reimbursements would be affected by implementing such a pricing schedule. We repeat this using expenditure-neutral NLA-based relative values compared with current submitted charges and discuss how reimbursements would be affected by implementing such a pricing schedule.

In summary, our scope of work is limited to analyzing charge data, developing charge-based relative values, and providing comparisons across several hypothetical fee schedules. We explore how sensitive the relative values are to various partitions of the charge data and methodologies for numeraire construction, and we provide comparisons to show how fees and revenues would change if a new price schedule were to be based on charge-based relative values. We analyze patterns that exist in current charge data, and in segmenting the analysis by various classes we are able to compare the effects of one approach over another. We do not attempt to predict any behavioral changes that might occur if a charge-based system were to be implemented; our analysis implicitly holds behavioral response constant when we conduct comparative analyses. Nor do we consider how behavioral responses might be mitigated by any transition or phase-in of any new system. These analyses are meant to inform CMS about properties of the existing charge data, as a point of departure for possible future development of a charge-based RVS.

2. JUSTIFICATION

Ideally, the new RVS to be used in the Medicare laboratory fee schedule would accurately reflect the relative costs in an efficient production process of different procedures (as defined by CMS's Healthcare Common Procedure Coding System [HCPCS] codes). If procedure A costs twice as much to produce as procedure B, then procedure A's relative value should be twice as high as procedure B's relative value. The rationale for using charges to determine relative values is that charges may be systematically related to costs.

We know that prices and submitted charges are consistently higher than the net prices actually received by laboratories (i.e., the amount laboratories are actually paid after subtracting contractual discounts and other allowances). Thus, under the reasonable assumption that net prices are greater than or equal to the costs of each laboratory procedure, it is clear that we cannot claim that submitted charges provide an accurate measure of costs. However, if prices are marked up over costs by the same percentage for all procedures, relative prices will provide an accurate measure of relative costs because the markup factor will cancel out of both prices. For example, suppose that the markup rate over cost is 80 percent for both procedures A and B, which have costs C_A and C_B and prices P_A and P_B, respectively. Then $P_A = 1.8\ C_A$, $P_B = 1.8\ C_B$, and the relative price of A with respect to B is equal to the relative costs of A and B:

$$P_A / P_B = C_A / C_B$$

Using standard economic theory, it is possible to generate relative prices that accurately reflect relative costs. Assuming firms maximize profits, they will set prices at the point where their marginal revenue equals marginal cost. This equation can be written as

$$P_1\ (1 - 1/E_1) = MC_1, \tag{2.1}$$

where P_1 is the price of procedure 1, E_1 is the elasticity of demand for procedure 1, and MC_1 is the marginal cost for procedure 1. If we divide the equations for procedures 1 and 2, we get

$$\frac{P_1\ (1 - 1/E_1)}{P_2\ (1 - 1/E_2)} = \frac{MC_1}{MC_2}\ . \qquad (2.2)$$

If the elasticity is the same for both procedures, the equation simplifies to the following:

$$P_1\ /\ P_2 = MC_1\ /\ MC_2 \qquad (2.3)$$

In this case, relative prices accurately reflect relative costs.

The preceding analysis relies on a number of assumptions. The result that relative prices reflect relative costs may not hold if the assumptions are violated. These violations could occur in several ways. First, elasticities may differ across different HCPCS codes. This may be true for newer procedures that are more reliable, thus exhibiting less elastic demand (and higher markups).[2] Second, the marginal costs may themselves be a function of the extant pricing schedules, to the extent that price distortions in the fee schedule distort technological investment and diffusion. Third, firms may not act rationally to maximize profits. Or, in consideration of information and transactions costs, it may be rational for them to adopt a single, universal markup over all procedures, rather than attempt to optimize the markup on each procedure. Fourth, complementarities or substitution effects may confound the relative costs and elasticities, especially if the profit maximization strategy is a more complicated product bundling scheme. Fifth, marginal costs for the same HCPCS codes may vary across firms for a variety of reasons, including scale of operations. Finally, laboratories may not set their charges (list prices) in a systematic way, because these charges are seldom used to set reimbursement rates. Or, they may set them according to an observed industry average, rather than as a reflection of their own production costs.

Despite the potential for violations of the assumptions, we believe that analyzing charges is still a useful starting point in the development of a consensus-based RVS. The assumption that

---

[2]On the other hand, because of the information and transaction costs of understanding individual product elasticities, firms might simply adopt a single, universal markup over costs for all procedures, based on an average elasticity across products, rather than attempt to optimally set the markup based on the elasticity of each individual product. This would support our general approach, because the common elasticity would cancel out of Equation 2.2.

laboratories base charges in part on costs is reasonable because even if laboratories do not know the per unit costs for procedures, they most likely have a knowledge of the relative costs. Based on these assumptions, we can provide analysis of relative values based on charges, present these findings to industry experts, and see if they have face validity. This analysis is useful as a starting point for further investigation, which can be directed by specific anomalies (or lack thereof) found in our analysis.

In summary, the charge-based approach has advantages and disadvantages:

Advantages

- Charge data are readily available. Under the current reimbursement system, laboratories submit charges to Medicare, and these charges are captured electronically. In addition, laboratories have price schedules that they use to charge their customers. In contrast, laboratories do not submit cost information and do not generally compile information on the cost of individual procedures. For a number of reasons, developing a microcosting approach could be difficult. Demonstration projects to negotiate fee schedules or conduct competitive bidding would also have to be designed and implemented. None of these approaches could be accomplished as quickly as performing an analysis of charge data.
- Charge data are current, while the basic information behind most of the current fee schedule prices dates from 1984.
- The charge-based relative values might identify procedures that are most likely to be over- or under-priced under an NLA-based payment system. Identifying these procedures could allow CMS or an expert panel to make adjustments in fees in the interim payment system, based on NLAs, prior to adoption of the long-term rational fee schedule envisioned by the IOM.
- Charge-based relative values could provide a starting point for a negotiated or consensus approach for setting an RVS. The consensus approach could be facilitated if the charge-based relative values generally have face validity.
- A charge-based NLA could also be used as a benchmark for evaluating other approaches.
- Potentially, charge-based weights can be updated over time as submitted charges change. Submitted charges might also be incorporated in the process of setting fees for new procedures.

Disadvantages

- Relative charges may not reflect relative costs if markups over cost differ between procedures. As noted above, the fact that submitted charges bear little resemblance to net prices will not pose a problem for forming an RVS if submitted charges are consistently discounted by the same percentage. If markups differ between procedures, the systematic relationship between costs and charges will be attenuated. Unfortunately, there is little evidence on whether percentage markups are constant across procedures.
- Laboratories may have incentives to distort future charge levels if the charges are used to set a fee schedule. However, this incentive will be diluted if charges are only used to set the RVS and not the conversion factor and by the fact that each laboratory has a relatively small impact on the overall distribution of relative charges.
- Because of data limitations, not all laboratories are considered in this study. The 2000 Physician/Supplier Procedure Summary Master File used here does not include data from hospital outpatient departments and other non-inpatient facilities. Therefore, the results of this study may not be generalized to all Part B laboratory providers. Still, included laboratories account for about 60 percent of Part B laboratory allowed charges (IOM, 2000).

As noted above, one of the disadvantages of the charge-based approach is its reliance on the assumption that prices are marked up over costs by the same percentage for all HCPCS codes. An alternative assumption is that the percentage markups are approximately equal for certain groups of HCPCS codes—for example, HCPCS codes in a product class, such as Hematology or Microbiology; HCPCS codes conducted within the same geographic regions; or HCPCS codes conducted at comparable production sites (e.g., physician laboratories or independent laboratories). Relative values can be calculated for different groups to investigate whether these group-wise measures vary significantly from more universal measures.

Weighing the advantages and disadvantages, we investigated the charge-based approach because charge data are readily available, and it is reasonable to assume that charges for individual procedures are correlated with procedure costs. At the same time, the disadvantages cannot readily be dismissed. Calculating separate charges for procedures grouped by type, region, and production platform can ameliorate the possibility that markups differ between procedures.

We believe that the charge-based approach can contribute to the development of a rational RVS for pricing purposes. But given the aforementioned difficulties, we recommend that charge data be used as one of the sources in the development of the RVS but not the only source. Charge data could be used as the starting point for a consensus panel of experts, who can make adjustments based on their expertise or experience, or as a benchmark for comparing the results of other approaches for developing the RVS.

3. METHODOLOGY

3.1 Data

We analyze the most recent available (2000) Physician/Supplier Procedure Summary Master file. The file includes information from all Part B claims submitted to Medicare carriers for reimbursement. The Physician/Supplier Procedure Summary data have an important limitation: only Part B claims submitted to Medicare carriers by independent laboratories and physician office laboratories are included. Part B claims submitted by hospital outpatient laboratories are not included in the data, because these claims are submitted to Medicare fiscal intermediaries. Unfortunately, hospital outpatient laboratory submitted charges are not available. Independent laboratories and physician office laboratories accounted for about 58 percent of Medicare Part B allowed charges in 1998 (IOM, 2000).

The data are aggregated at the HCPCS code, physician supplier specialty code, carrier number, locality code, region code, service type, and place of service level. To conduct the main analysis (Section 4.1), we further aggregate the submitted charges (net of denied charges) and total services count (net of denied services count) to the HCPCS code level. We only examine HCPCS codes that are used in the Clinical Diagnostic Laboratory Fee Schedule and have NLAs.[3] Initially, we include all laboratories in the analysis. We describe subanalyses below that are performed on subsets of laboratories. In one subanalysis, we compare physician office laboratories and independent laboratories. To perform this analysis, we aggregate the data to the HCPCS code by type of laboratory. In another subanalysis, we examine regional variation and aggregate the data to the HCPCS code at the regional level and compare each region to the national level.

Because we want to show how reimbursements would shift if a new charge-based pricing system were to be implemented, we use as the comparison the current NLA-based pricing system. This is actually a blend of NLAs and regional fee schedule amounts because in the current system, reimbursement equals the lower of the NLA amount, the national fee schedule amount, or the submitted charge. Thus, our "NLA-based" relative values are actually based on whichever fee schedule is binding, which is heavily dominated by the NLAs. These fee schedules are posted annually in the Clinical Diagnostic Laboratory Fee Schedule.

Before performing the analysis, we compared prices implied by the allowed charges and volume to the actual NLA-based reimbursement prices. This comparison allowed us to check for very large discrepancies, which would suggest problems with the data. Upon further investigation of some very large discrepancies, we determined that the NLA fee schedule amount for a set of nine procedures (88150, 88152, 88153, 88154, 88164, 88165, 88166, 88167, and P3000) had been updated by CMS after our edition of the fee schedule was published. Using these updates, we adjusted the NLA-based amount for the remaining analysis.[4]

Inclusion in a panel explained many of the remaining discrepancies between prices implied by the allowed charges and

---

[3]Our analysis did not include G0001, drawing blood for specimen in our analysis, because there is no official NLA for this procedure in the fee schedule. However, the fee is set at $3.00 in every state. After performing our main analysis, we performed a secondary analysis to calculate the relative value and expenditure-neutral price for G0001, based on the procedure's average submitted charge. The expenditure-neutral price based on the relative value would be $3.02.

[4]For example, the price for HCPCS code P3000 changed after the NLA was published. In the IOM (2000) report, we found the following: "In the 1999 Balanced Budget Reconciliation Act (BBRA), Congress doubled the minimum payment for Pap tests.[11]" Footnote 11 states: "Raising the price from $7.15 to $14.60" (p. 91).

volume and the actual NLA-based prices. About 5.34 percent of all HCPCS codes were determined to be either an approved panel procedure or one of the 22 individual procedures that are reimbursed as an automated test panel (ATP).[5] The CMS-approved panel procedures make up the test class "Panels" in this report, and the procedures within them are often, but not always, components of ATPs. Approved panels must contain a specific minimum set of procedures, and they are reimbursed at the lesser of the panel procedure HCPCS rate or the sum of the rates for the individual HCPCS procedure components in the panel. In contrast, reimbursement for an ATP is based on the number of individual test components, rather than the individual procedures themselves. For example, ATP03 represents three automated tests, regardless of the identity of the individual tests that make up the panel. This poses problems for the analysis because, in our data set, allowed charges are allocated to the HCPCS codes of the individual test components, whereas the recorded reimbursement for these procedures depends on whether a given procedure was submitted alone or as part of a panel. If the procedure was submitted as part of a panel, the reimbursement for the procedure is based on the reimbursement rate for the entire panel. Specifically, the panel reimbursement is allocated to each individual procedure based on the share of submitted charges for that procedure.[6]

For example, consider a panel of two procedures (i.e., ATP02) that is reimbursed at $10. If the submitted charge for procedure A is $15 and the submitted charge for procedure B is $5, then the allowed charge for procedure A would be $7.50 = $10 * (15/20). Similarly, the allowed charge for procedure B would be $2.50 = $10 * (5/20). However, if a procedure was submitted alone, the reimbursement rate is often (although not always) the NLA.

Most of the implied allowed charges for HCPCS codes of panels, panel components, or ATP procedures were much lower than the

---

[5]The 22 procedures processed as ATPs are listed in the HCFA Publication in November 1999, Program Memorandum (PM) Transmittal AB-99-85.

[6]See HCFA Publication, 60AB: Program Memorandum (PM) Transmittal AB-97-5 issued in March 1997. This memorandum describes the process for adjudicating new laboratory panels approved by the American Medical Association Current Procedural Terminology (CPT) Board at their November 1996 meeting. The new CPT codes must be used for service dates January 1, 1998 and later.

NLA for the HCPCS. To facilitate comparison of the charge-based relative values to the NLAs, we created the following adjusted volume variable for each procedure i:

$$\text{Adjusted Volume} = \text{Total Allowed Charges}_i \ / \ \text{NLA-Based Price}_i$$

The adjusted volume shows how many units of the procedure would be reimbursed if each unit was reimbursed at the NLA rate. This adjustment only affects the weighting of each procedure when we sum to the national, regional, or laboratory-type level. We determined that the remaining discrepancies were small enough to be artifacts of the system wherein CMS reimburses laboratories at the minimum of the submitted charge and the allowed charge. Since it is not possible to accurately predict the extent to which laboratories submit charges that are less than allowed charges, we needed to adjust the relative values of these procedures, similar to adjustments made for the panel procedures.

3.2 Derivation of Relative Charges

In the payment system envisioned by the IOM, payments for a HCPCS code would equal the relative value for the code times a universal conversion factor that applies to all codes. The relative values are measured relative to a unit of measure. The conversion factor, equal to $X/relative value unit, converts the relative values into dollars.

To transform the charge data into relative values, we divide each charge by a standard amount, or numeraire. As described below, in our analysis, we use the weighted average charge as the numeraire. The numeraire itself has a relative value of one (equal to the weighted average charge divided by the weighted average charge); hence the term numeraire. The relative value of each procedure will then be expressed as a number that is either greater or less than one.

To calculate the average submitted charge for each procedure (HCPCS code), we divide the total submitted charges by the total services count. Next, we divide the average charge for each

HCPCS code by the numeraire.[7] Given the invariance of relative values to the choice of numeraire, it is desirable to choose a numeraire that is easy to interpret. Possibilities include a weighted average of the submitted charges (with weights based on the adjusted volume of each procedure) or the submitted charge for a common procedure. The main advantage of using the weighted average charge as a numeraire is that we can readily interpret the relative charge for an individual procedure as being proportionally lower or higher than the average charge for all procedures. Alternatively, the advantage of using the charge for a common procedure as the numeraire is that the common procedure may provide a useful benchmark for industry experts in assessing relative values.

Ultimately, we compare relative values based on submitted charges to relative values that are generated from NLAs. For this comparison, the choice of numeraire *does* matter, in a way that favors the use of the weighted average prices as the numeraire. Different RVSs can be generated using submitted charges or NLAs. If the scales have the same conversion factor, we will be able to directly compare the relative value units across the different scales. It can be shown that the conversion factors associated with each of the scales will be the same if the numeraires are based on the average submitted charge and average NLA, respectively. If the numeraires were instead based on a common procedure's submitted charge or NLA, the conversion factors associated with each scale would be different, and we would be unable to directly compare the relative values across scales. Therefore, we use the average charge for all procedures as the numeraire for the charge-based RVS, and the average NLA as the numeraire for the NLA-based RVS.

In addition to the main analysis, which uses data for all HCPCS codes, we also perform separate analyses for each major class of

[7] Ease of interpretation and comparability between RVSs are the primary considerations when choosing a numeraire. Within a single RVS, relative values are invariant to the choice of numeraire. For example, suppose that procedure A has a price of 1 and procedure B has a price of 2. If A is chosen as the numeraire, A will be associated with 1 relative value unit and B will be associated with 2 relative value units. Conversely, if B is chosen as the numeraire, A will be associated with 0.5 relative value units and B will be associated with 1 relative value unit. In both cases, there will be 1:2 ratio between A's and B's relative values.

HCPCS codes. We have identified 12 classes: Organ and Disease Panels (80048–80090), Drug Testing (80100–80103), Therapeutic Drug Assays (80150–80299), Urinalysis (81000–81099), Chemistry and Toxicology (82000–84999, G0103, G0107), Hematology (85000–85999), Immunology (86000–86849), Transfusion Medicine (86850–86999), Microbiology (87001–87999, Q0111, Q0114), Cytopathology (88130–88199; P3000), Cytogenetic Studies (88230–88299), and Other (78267–78268, 88371–88372, 89050–89399, Q0115, G0026–G0027, Q0112–Q0113).

The 12 classes vary in the numbers of different procedures (HCPCS codes) included in them. Chemistry and Toxicology is the largest class, with 416 procedures within it (e.g., tests for amino acids, minerals, carbon dioxide, and cholesterol). Immunology includes 154 procedures (e.g., tests for allergens or antibodies), and Microbiology includes 152 procedures (e.g., tests for the presence of organisms or colonies of organisms, fungus, or parasites). Hematology is the next largest class, with 95 procedures (e.g., blood platelet count, white cell count, clot reaction, and clotting inhibitors). The remaining 8 classes have 30 or fewer procedures within them (see Table 4-1, which lists the number of different procedures within each class).

Class size and volume of procedures may not be directly related. Test volume is the total number of tests done for each procedure type. The Hematology class has 95 different procedures within it, accounting for about 10 percent of all procedure types but accounting for over 25 percent of the total volume of tests done in the laboratories studied here. The Immunology class, by comparison, accounts for about 17 percent of all procedure types but only about 3 percent of the total volume of tests performed. The largest class, Chemistry and Toxicology, accounts for about 44 percent of the different procedures and about 40 percent of the total volume of tests.

In class-specific analyses, we calculate relative values and numeraires for the submitted charge data by class, using only data from each individual class. We calculate relative values for the NLAs in a similar fashion. For example, the numeraire for an NLA-based relative value for a HCPCS code in class “A” is the volume-weighted average of the NLAs for all HCPCS codes performed within this particular class.

One reason for performing class-specific analyses is that demand elasticities may vary less within some classes than they do across all classes. As Eq. (2.2) indicates, relative prices will better reflect relative costs if the elasticities of demand are similar across procedures.

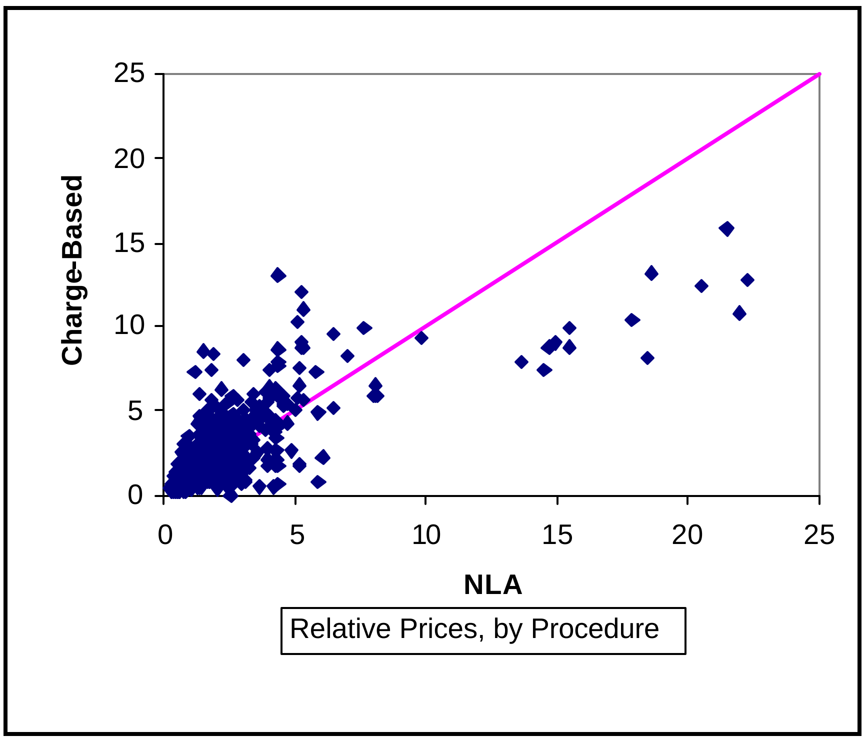
4. RESULTS FROM THE ANALYSIS

4.1 Comparison of Charge-Based Relative Values to NLA-Based Relative Values

4.1.1 Main Analysis, with Universal Numeraire Based on All Procedures

First, we graph the relative values based on submitted charges, for each procedure, against their relative values based on the NLA (Figure 4-1). Each point on the scatterplot represents a pair of values: the charge-based relative value for the procedure, and the NLA-based relative value for the procedure. The plot shows

Figure 4-1. NLA-based versus Charge-based Relative Values, by Procedure, with Universal Numeraire



considerable dispersion from the hypothetical straight line, suggesting that the two value schedules are different. The vast majority of the procedures have relative values less than 5. There are two sets of extreme outliers. One set of procedures has NLA-based relative values greater than 15 but charge-based relative values less than 15; these procedures appear on the right-hand side of the chart below the 45-degree line. The procedures are all in the Cytogenetic Studies class. As we show and explain below, these outliers result in part from using a universal numeraire. Another set of outliers are the procedures with charge-based relative values greater than 7 and NLA-based relative values generally less than 7. These procedures are in the mid-left-hand side of the figure above the 45-degree line. The majority of these procedures are in the Microbiology class. As we show below, these outliers are not a result of using a universal numeraire—they persist even when a class-based numeraire is used.

Overall, the scatterplot suggests considerable differences between the submitted charge-based relative values and the NLA-based relative values. Next, we analyze these differences statistically to see whether the observed differences are statistically significant. A statistically significant test statistic (one with a p-value less than or equal to 0.05) suggests that the differences in the two distributions are large enough to be unlikely to have occurred due to random chance. With a 0.05 p-value cutoff, there is a 5 percent significance level associated with the test, which means that we are 95 percent sure that the observed result could not have occurred by chance, when the distributions were actually similar. Thus, we can be 95 percent confident with our conclusion for the hypothesis test. When the statistical test produces a significant result, we conclude that the two distributions are significantly different.

For this statistical comparison of the charge-based and the NLA-based relative value schedules, we perform a Kolmogorov–Smirnov equality of distributions test (described in Appendix A). The test allows us to determine whether this pair of relative value schedules is distributed in a similar manner. A significant test statistic (low p-value) would suggest that the distributions are not the same, so that for at least some HCPCS codes, pricing would be quite different under the two systems compared. These results are presented in Table 4-1. With a p-value of ≤ 0.01, the Kolmogorov–Smirnov test suggests that we can be at least 99 percent confident that the two distributions are different. If the two distributions are not significantly different, we can conclude that an NLA-based relative value is very similar to a charge-based relative value.

The ultimate purpose of conducting this test is to see whether switching from the current NLA-based system to a charge-based system would result in significant changes in reimbursements across procedures. When the test is not significant, we conclude that the two relative value schedules are not different. Because the relative values are simply a standardized expression of relative prices, we can also conclude that the relative prices implied by these relative values are not significantly different. In this case, because the NLA-based pricing system is already in use, there would be no impact from changing to a charge-based pricing system (and hence no reason to change, since they are so similar). When the two schedules are significantly different, we know that changing to the charge-based system could result in significant changes in the distribution of payments. Unfortunately, in this case the test does not tell us which schedule is the "correct" one to use, only that they are significantly different.

Table 4-1. Summary of Relative Values with Universal Numeraire Based on All Procedures

	Number of Different Procedures[a]	Kolmogorov-Smirnov Test (p-value)	Proportion of Procedures with Submitted (Charge-based) Relative Values that are:		Proportion of Procedures that Are Outliers[b]	
			> NLA-based	< NLA-based	> NLA-based	< NLA-based
All Procedures	936	0.00	57%	42%	12%	10%
Class:						
Chemistry and Toxicology	416	0.04	55%	44%	10%	12%
Cytogenetic Studies	24	0.00	21%	79%	4%	21%
Cytopathology	12	0.00	8%	92%	8%	17%
Drug Testing	3	0.49	33%	67%	0%	33%
Hematology	95	0.00	76%	23%	37%	0%
Immunology	154	0.00	64%	34%	8%	5%
Microbiology	152	0.05	49%	51%	7%	16%
Panels	7	0.90	43%	43%	0%	0%
Therapeutic Drug Assays	30	0.00	87%	13%	13%	3%
Transfusion Medicine	10	0.29	60%	40%	10%	0%
Urinalysis	10	0.98	50%	50%	0%	0%
Other	23	0.55	61%	39%	13%	4%

[a]Number of different procedures with nonzero values in the NLA fee schedule.

[b]Outliers are defined as those procedures that exhibit percentage differences (between the charge-based relative value and the NLA-based relative value) that are more than one standard deviation from the mean percentage difference.

To facilitate comparison with Section 4.1.3, results in Table 4-1 and Appendix Table B-1 are also presented by procedure class. The Kolmogorov–Smirnov test is recalculated in Table 4-1 for each of the 12 classes to determine whether differences in the two schedules are more pronounced in some classes than in others. We find that all of the classes that contain 30 or more separate procedures in them, and 2 of the 7 smaller classes, lead to the same conclusion as the main analysis test: the two distributions are different. Five of the seven smaller classes have large p-values, and the hypothesis of equal distributions cannot be rejected.[8] In practical terms, these findings suggest that for all procedures taken as a group, and also for the seven largest classes taken individually, prices based on the charge-based relative values would be significantly different from the current NLA-based system. We calculate class-specific numeraires in Section 4.1.2 and repeat this analysis to examine whether this change (numeraire-methodology change) in computation of the relative values results in less discrepancy between the two schedules.

Columns 3 and 4 in Table 4-1 list the proportion of procedures where the submitted charge-based relative values are greater than or less than the NLA-based relative values. If the submitted charge-based relative value is greater than the NLA-based relative value, it implies that the price for that procedure would increase if the submitted charge-based relative value was used. We find that 55 percent of the procedures in the Chemistry and Toxicology class would have relatively high prices using submitted charge-based relative values, whereas 44 percent of the procedures would have relatively low prices. It is clear from the table that, assuming laboratories do not change their pricing behavior in response to a new pricing system, the Cytogenetic Studies and Cytopathology classes would have, for the majority of the procedures in the class, relatively lower prices under a charge-based RVS with a universal numeraire. In contrast, the Immunology and Therapeutic Drug Assays classes would have relatively higher prices under a charge-based RVS with a universal numeraire.

We define outlier procedures as those with unusually large percentage differences relative to the mean percentage difference. In our data, the mean percentage difference (between charge-based and NLA-based relative values) over all procedures combined is 19 percent; that is, on average, the charge-based relative values are 19 percent higher than the NLA-based relative

[8] The Kolmogorov–Smirnov test (a distribution-free test) generally has lower statistical power to discriminate between a true and a false null than distribution-dependent tests, like the t-test. We cannot be sure here that failure to reject the hypothesis of similar distributions is not due to low power rather than due to similar distributions.

values.[9] For an absolutely symmetrical distribution of percentage differences, we would expect a mean of zero—half would be negative, and half would be positive. So our distribution of percentage differences is somewhat skewed. We define an outlier as a procedure with a percentage difference greater than one standard deviation from the mean percentage difference. One standard deviation from the mean percentage difference is ±63 percent. Therefore, percentage differences that are larger than 82 percent (19 percent + 63 percent) or smaller than –44 percent (19 percent—63 percent) are designated as outliers. About 22 percent of the sample (206 procedures) qualify as outliers by this definition. The number of outliers is not important (for example, in a normal distribution, 31.7 percent of observations would be expected to be outliers, based on a one standard deviation difference). What is important is the large magnitude of the standard deviation and its implication for prices if charge-based relative values were to replace NLA-based relative values.

Because the standard deviation is large (a 63 percent difference), prices for the outlier procedures would either decrease by at least 44 percent or increase by at least 82 percent if charge-based relative values were adopted in place of NLA-based relative values (assuming a constant conversion factor). These represent large price changes. With changes this large, a gradual phase-in might be desirable, if the new system were adopted, to allow laboratories time to adjust to the new prices.

The last two columns of Table 4-1 show the proportion of procedures that are outliers. We consider the percentage of outliers to be "large" if they exceed what would occur under a normal distribution, which is 31.7 percent. By this definition of "large," Drug Testing has a large percentage of outliers that are less than the NLA-based relative price, whereas Hematology has a large percentage of outliers that are more than the NLA-based relative price. In Section 4.1.2, we discuss whether these outliers

[9] Procedures with large positive differences tend to have low volume, and procedures with large negative differences tend to have high volume. As a result, the mean percentage difference across procedures (which is not weighted for procedure volume) is positive, while the average relative charge (weighted for volume) is 1 for both the charge-based and the NLA-based systems.

seem to be an artifact of the universal numeraire or a reflection of actual differences between relative costs of these procedures.

Next in our comparison of charge-based relative values to NLA-based relative values, we list outlier procedures, by HCPCS code and class (Appendix B, Table B-1). A positive percentage difference means that the charge-based relative value exceeds the NLA-based relative value, while a negative percentage difference means that the charge-based relative value is less than the NLA-based relative value. For example, procedure 82000 (assay of blood acetaldehyde) under Chemistry and Toxicology has a difference of 218 percent. This means that the charge-based relative value exceeds the NLA-based relative value by about 218 percent for this procedure. In practical terms, if this charge-based relative value schedule were used to replace the current NLA-based system, the fee to this procedure would rise about 218 percentage points, assuming a constant conversion factor, no behavior changes, and no changes in overall laboratory allowed charges. (The charge-based relative value is 4.45, which is 218 percent higher than the NLA-based relative value of 1.4.) The table shows that, for some classes (like Chemistry and Toxicology), some of the included procedures would gain and others would lose with this sort of change in reimbursement. For others, like Hematology, all of the outlier procedures listed in Table B-1 would gain with this change in reimbursement.

4.1.2 Subanalyses, with Numeraire Based on Procedure Class Subset

The main reason for performing this subanalysis is because differences in the standard markups (of charge above actual cost) across classes may be due to underlying differences in the elasticities of demand for the classes of procedures. When the standard markups are different across classes, their relative values should be based on class-specific numeraires. We are also interested in determining whether the choice of numeraire contributes to the differences in the outlier activity noted when we compare Tables 4-1 and 4-2. It could be that defining the numeraire based on the universe of all procedures, rather than defining it separately by class, introduces noise into the relative value comparisons that is not due to differences in standard

Table 4-2. Summary of Relative Values with Numeraire Based on Class

	Number of Procedures[a]	Kolmogorov –Smirnov Test (p-value)	Proportion of Procedures with Submitted (Charge-based) Relative Values that are:		Proportion of Procedures that Are Outliers[b]	
			> NLA-based	< NLA-based	> NLA-based	< NLA-based
All Procedures	936	0.02	56%	43%	12%	9%
Class:						
Chemistry and Toxicology	416	0.03	56%	43%	10%	12%
Cytogenetic Studies	24	0.58	54%	42%	21%	0%
Cytopathology	12	0.06	25%	75%	8%	8%
Drug Testing	3	0.49	67%	33%	0%	0%
Hematology	95	0.00	79%	21%	41%	0%
Immunology	154	0.05	51%	47%	7%	6%
Microbiology	152	0.03	48%	51%	7%	16%
Panels	7	0.90	57%	43%	0%	0%
Therapeutic Drug Assays	30	0.02	70%	30%	3%	3%
Transfusion Medicine	10	0.29	20%	80%	10%	0%
Urinalysis	10	0.29	30%	70%	0%	0%
Other	23	0.98	61%	39%	9%	9%

[a]Number of procedures with nonzero values in the NLA fee schedule.

[b]Outliers are defined by class, as those procedures that have a percentage difference (between the charge-based relative value and the NLA-based relative value) that is more than one standard deviation from the mean percentage difference for the class.

markups across classes. So we repeat the above analysis (Section 4.1.1) using the class-based numeraires, but using the same bounds for outlier definition (percentage differences exceeding 0.19±0.63). With this refinement to the numeraire methodology, we hope to get a cleaner picture regarding differences in the value scales that may be due to differences in markups.

First, we graph the charge-based relative values against the NLA-based relative values, for each class (Appendix C, Figures C-1 through C-12). The plots show considerable dispersion from the hypothetical straight line in the majority of classes, suggesting that the two value schedules are different. We see from the plots that the charge-based relative values appear to be higher than the NLA-based relative values for the majority of procedures, and especially for the Chemistry and Toxicology, Hematology, and Therapeutic Drug Assays classes. Note that while the Cytogenetic Studies class appeared to have many outliers when the universal numeraire was used (see Table 4-1, Table B-1, and the cluster of outliers identified in Figure 4-1), the relative values based on class now appear to be very close to each other (Figure C-2). However the dispersion seen when using the universal numeraire apparently remains for most of the other classes when using the class-based numeraire (Figures C-1, C-3 through C-8, and C-10 through C-12).

The scatterplots suggest considerable differences between the charge-based relative values and the NLA-based relative values. Next, we analyze these differences statistically to see whether the observed differences are statistically significant. We compare the relative values calculated separately for each class of HCPCS code to those calculated similarly from the NLA, and conduct the Kolmogorov–Smirnov equality of distributions test for each class. These results are presented in Table 4-2.

The overall Kolmogorov–Smirnov test statistic has a p-value of 0.02, which suggests that we can be about 98 percent confident that the two distributions of relative values are significantly different. We find that those classes that contain 30 or more separate procedures in them lead to the same conclusion as the overall test: that the two distributions are different. For all seven classes with fewer components, the p-values are larger and the hypothesis of equal distributions cannot be rejected.

In practical terms, these findings suggest that, for all procedures taken as a group, and also for the five largest classes taken individually (Chemistry and Toxicology, Hematology, Immunology, Microbiology, and Therapeutic Drug Assays), prices based on the charge-based relative values would be significantly different from the current NLA-based system. This is the same result we found when we compared the charge-based relative values with a universal numeraire to the current NLA-based relative values.

In Table 4-2, we show a summary of the relative values, with numeraire based on class. This table can be directly compared with Table 4-1, the summary of relative values with universal numeraire. To facilitate this comparison, we defined outliers using the same parameters as in Table 4-1: the charge-based relative value is an outlier if it is less than 44 percent lower or more than 88 percent higher than the procedure's NLA-based relative value.

In the last two columns of Table 4-2, we show the percentage of outliers for each class. As in Table 4-1, we consider the percentage of outliers to be "large" if they exceed what would occur under a normal distribution, which is 31.7 percent. By this definition of "large," Hematology has a large percentage of outliers, as it did in Table 4-1. So, the class-based numeraire methodology did not significantly reduce the volume of outliers for this class (as compared to the universal numeraire). Counting the number of outliers (not shown in the table), we find that there are only 9 fewer outliers over all classes combined, when using the class-based numeraire instead of the universal numeraire. In comparison to Table 4-1, the percentage of outliers in the last two columns has an unchanged distribution for two of the largest classes (Chemistry and Toxicology, Microbiology). This means that the outliers in these classes are not due to the numeraire methodology—they persist whichever method is used.

When the class-based numeraires are used, the distribution of outliers becomes more symmetrical for Immunology, Cytopathology, Drug Testing, and Therapeutic Drug Assays. For example, in Table 4-1, the Cytopathology class had one (8 percent) of its procedures defined as outliers with charged-based relative values greater than NLA-based, and two (17 percent) with values less than the NLA. When the class-based numeraire is used, there is one less outlier in this class overall, and the two remaining outliers are evenly dispersed, with one (8 percent) of them greater than the NLA-based relative values and one (8 percent) of them less than the NLA-based relative values. This reflects a property of class-based numeraires: increases and decreases are dispersed within classes rather than between classes.

In one class—Cytogenetic Studies (identified earlier as the class with the most problematic outlier problem under the universal numeraire methodology, the cluster of outliers in Figure C-2)—the large proportion of outliers persists after the numeraire is changed. But the distribution of the outliers shifts with the change in numeraire—the proportion that exceeds the NLA rises (from 4 percent to 21 percent) while the proportion that is less than the NLA falls (from 21 percent to 0 percent). With the class-based numeraire, this class has one less outlier, and the pattern for the five that remain has shifted from being all below the NLA to all above it. This shift reflects the fact that this small group of procedures is among the most expensive, with many NLAs exceeding $200 per test. The universal numeraire pushes down their charge-based relative values relative to the class-based method because the universal average (based on all procedures) is so much lower than the class-specific average.

Next, we list those procedures that are outliers when the charge-based relative values are calculated using class-specific numeraires. We list these procedures by HCPCS code and class in Appendix Table B-2. A positive percentage difference means that the charge-based relative value exceeds the NLA-based relative value, while a negative percentage difference means that the charge-based relative value is less than the NLA-based relative value. For example, procedure 82000 (assay of blood acetaldehyde) under Chemistry and Toxicology has a difference of 225. This means that the charge-based relative value exceeds the NLA-based relative value by about 225 percent for this procedure. In practical terms, if this charge-based relative value schedule were to replace the current NLA-based system, the fee for this procedure would ultimately rise about 225 percentage points, assuming a constant conversion factor, no behavior changes, and no changes in overall laboratory allowed charges. Table B-2 (using a class-based numeraire) is consistent with Table B-1 (using a universal numeraire): for all but one class, some of the outlier procedures listed there would gain and others would lose with this sort of change in reimbursement. Hematology would unambiguously gain across all of the listed outlier procedures (Tables B-1 and B-2).

4.1.3 Implications of Different Numeraires

In this section, we compare the results from Sections 4.1.1 and 4.1.2. In Sections 4.1.1 and 4.1.2, we did separate analyses each with a unique assumption about the methodology for calculating the numeraire (in Section 4.1.1, using the entire sample, or universe; in Section 4.1.2, using subsets, by procedure class). In all cases, we used the volume-weighted average of the respective prices as the numeraire.

In Table 4-3, we summarize the effects of numeraire choice by comparing the charge-based and the NLA-based relative value distributions. This table succinctly summarizes the differences in the proportions found in columns 3 through 6 in Tables 4-1 (universal numeraire) and 4-2 (class-based numeraire). The first column shows, by procedure class, the proportion of the procedures with charge-based relative values exceeding the NLA-based relative values, for those situations where the choice of numeraire does not matter (either numeraire results in the same ordering). Thus, for Chemistry and Toxicology, 56 percent of the procedures in this class have charge-based relative values exceeding NLA-based relative values, irrespective of which numeraire is chosen. The fourth column also shows those procedures with agreement across numeraire choice in finding NLA-based relative values exceeding charge-based relative values.

The second and third columns show additional instances where the ordering of charge-based relative values to NLA-based relative values is sensitive to the numeraire chosen. That is, columns 2 and 3 show those situations where the two methods produce conflicting ordering of the relative values. These columns are of primary interest, because they help reveal which classes face increases or decreases in relative values depending on whether a universal or class-based numeraire is used.

Table 4-3 is interpreted as follows: for the first class of procedures, Chemistry and Toxicology, 56 percent of the procedures have charge-based relative values that are greater than or equal to NLA-based relative values, whether they are based on the universal or the class-specific numeraire. For those computed using the by-class numeraire method, the NLA-based relative value is exceeded an additional 1 percent of the time. The last column shows that irrespective of numeraire-construction methodology, procedures in this class agree 43 percent of the time regarding whether the submitted charge-based relative value is less than the NLA-based relative value.

For procedure classes Chemistry and Toxicology, Hematology, Microbiology, Panels, and Therapeutic Drug Assays, the choice of

Table 4-3. Summary of the Effect of the Numeraire Choice on Agreement between the Relative Value Distributions

		Proportion of Procedures with Charge-based Relative Values Greater than or Equal to NLA:			Proportion of Procedures with Charge-based Relative Values Less than NLA:
	Number of Procedures	**Charge-based Relative Values Using Either Numeraire Agree**	**Charge-based Relative Values Using the By-Class Numeraire**	**Charge-based Relative Values Using the Universal Numeraire**	**Charge-based Relative Values Using Either Numeraire Agree**
Chemistry and Toxicology	416	56%	1%	0%	43%
Cytogenetic Studies	24	21%	38%	0%	42%
Cytopathology	12	8%	17%	0%	75%
Drug Testing	3	33%	33%	0%	33%
Hematology	95	77%	2%	0%	21%
Immunology	154	53%	0%	13%	34%
Microbiology	152	49%	0%	1%	51%
Panels	7	61%	0%	0%	39%
Therapeutic Drug Assays	30	57%	0%	0%	43%
Transfusion Medicine	10	70%	0%	17%	13%
Urinalysis	10	20%	0%	40%	40%
Other	23	30%	0%	20%	50%

methodology for constructing the numeraire really does not impact the (direction of the) comparison of the charge-based with the

NLA-based relative values. For other procedures, using a class-specific numeraire increases the charge-based relative values above the NLA-based relative values for a significant proportion of additional procedures (Cytogenetic Studies, Cytopathology, Drug Testing). If a fee schedule were to be based on these by-class relative values, laboratories conducting these procedures (Cytogenetic Studies, Cytopathology, Drug Testing) would receive higher prices for more procedures than they would under a universal numeraire methodology. For several other classes, the opposite holds true: laboratories conducting these procedures would receive lower prices for more procedures if based on the class-based numeraire methodology (Immunology, Transfusion Medicine, Urinalysis).

In Appendix Table B-3, we list by procedure code those procedures with disagreement (between the relative size of the charge-based relative value and the NLA-based relative value) due to the numeraire methodology used. For example, under Chemistry and Toxicology, procedure 82671 (assay of estrogens), the difference between the charge-based and NLA-based relative value is negative using a universal numeraire and positive when using a class-based numeraire. This means that laboratories conducting this procedure would receive higher prices using the class-based numeraire but lower prices using the universal numeraire. We note that the percentage differences in Table B-3 are generally small. For most of the procedures in Table B-3, the charge-based relative value based on the universal numeraire is fairly close to the NLA-based relative value. It takes only a minor change with the individual class to shift the charge-based relative value to the opposite side of the NLA-based relative value. This is because the procedures listed in Table B-3 are not those with huge (outlier) differences. Therefore, we can conclude that where the two methods cause disagreement in the ranking of the charge-based relative to NLA-based relative values, the differences are small. The outlier cases tend to be in agreement (as regards ranking the two relative values) across the two methods of constructing the numeraire.

The main implication of using a class-based numeraire versus a universal numeraire is that the class-based numeraire facilitates calculation of expenditure neutral prices by class. In Section 5, we revisit the comparison of class-based and universal numeraires by comparing expenditure-neutral prices under each numeraire-construction methodology. This analysis reveals that there is some redistribution between classes if the universal numeraire is used. If class-based numeraires are used, no redistribution occurs. In the sensitivity analyses in the next section, we use class-based relative values, and compare relative values calculated over different subsets of the data.

4.2 Sensitivity Analyses

The purpose of the sensitivity analyses in this section is to test whether the relative values are robust to changes in the sample subsets over which they are calculated, in order to better understand variation in relative values. Some of this variation is expected due to noise in smaller samples (e.g., low volume versus high-volume procedures) or noise from uncertainty about how payments are recorded (panel tests). Some variation has natural policy implications. For example, it may be interesting to know whether independent and physician laboratories exhibit different significance levels (using the Kolmogorov–Smirnov test) for their deviations between the charge-based and the NLA-based relative value schedules. Also, we might expect differences in standard markups (of charges over costs) across procedure classes due to differences in average demand elasticities for the classes. In this case, the relative value schedules should be calculated separately for each class, rather than uniformly across all classes. Otherwise, we could introduce distortions in the comparisons of fee schedules that are due to invalid assumptions about these markups. In this section, we attempt to further explain differences in the charge-based versus the NLA-based relative values. Here we use various methods and compare various subsets of the data: by procedure volume, by laboratory type, and by region.

4.2.1 Procedure Volume

Many laboratory procedures are performed rarely. Because we rely on the law of large numbers to be confident that our average charge for a procedure is close to the true value in the underlying distribution, we expect that low volume procedures (defined as less than 500 test procedures conducted across all laboratories for a particular HCPCS code in a given year) may produce less reliable average submitted charge estimates. In turn, this could cause distortions in the charge-based relative values. These distortions could cause low-volume procedures to account for a disproportionate share of the outliers identified in previous analyses.

In this part of the sensitivity analysis, we compare the proportion of all procedures that are outliers to the proportion of low-volume procedures that are outliers (Table 4-4). We see that only 23 percent of all procedures are outliers, while a much larger proportion—34 percent—of low-volume procedures are outliers. Next, we look class by class and see the same pattern: generally, low-volume procedures are more likely to be outliers than all

Table 4-4. Summary of Relative Values—Analysis of the Effect of Volume and Inclusion in a Panel

	All Procedures		**Low-Volume Procedures**		**Panels and/or Procedures Processed as ATPs**	
	Number of Procedures[a]	**Proportion of Outliers[b]**	**Number of Procedures[a]**	**Proportion of Outliers[b]**	**Number of Procedures[a]**	**Proportion of Outliers[b]**
All Procedures	936	23%	301	34%	50	22%
Chemistry and Toxicology	416	24%	123	34%	26	27%
Cytogenetic Studies	24	4%	15	7%	0	0%
Cytopathology	12	8%	4	25%	0	0%
Drug Testing	3	33%	0	0%	0	0%
Hematology	95	32%	32	38%	3	0%
Immunology	154	13%	48	25%	12	0%
Microbiology	152	30%	63	51%	1	0%
Panels	7	57%	0	0%	7	57%
Therapeutic Drug Assays	30	27%	1	100%	0	0%
Transfusion Medicine	10	10%	3	0%	1	0%
Urinalysis	10	20%	0	0%	0	0%
Other	23	17%	12	8%	0	0%

[a]Number of procedures with nonzero prices in the NLA fee schedule.

[b]Outliers are defined as those procedures that are more than one standard deviation from the mean percentage difference between the charge-based and the NLA-based relative value.

procedures taken together. We find that low volume outliers tend to be large positive values, which means that the charge-based RVs are higher than the NLA-based RVs, and vice-versa.

4.2.2 Panels and/or Procedures Processed as ATPs

Individual procedures that are processed as parts of approved panels or automated multi-channel test panels (ATPs) may pose problems for charge-based relative value calculations. It is not clear whether submitted charges for these procedures are based on the cost of producing each test alone or on the cost of producing these in combination with other tests. It is possible that this uncertainty could result in these procedures having charge-based relative values that are less comparable to underlying costs (than non-panel related tests). Thus, these panel-related procedures might be more likely to appear as outliers. However, when panel tests are outliers, they tend to be negative outliers, meaning that the charge-based relative value is substantially less than the NLA-based relative value. Table 4-4 shows that the outlier rate for these panel procedures is about the same as the outlier rate across all procedures (22 percent versus 23 percent).

4.2.3 Laboratory Type

For policy purposes, it may be necessary to decide whether all laboratories should receive the same fees or whether fees should differ by type of laboratory (independent versus physician office). Relative charges may vary by type of laboratory. For example, procedures with significant economies of scale and scope might have lower charges at independent laboratories than physician office laboratories, and relative costs may differ between laboratory types. To shed light on the relative cost issue, first we plot the relative charges for each type of laboratory in Figures C-13 through C-25. This, and the following analyses, are limited to procedures that are produced in both independent laboratories and in physician office laboratories. The plots show substantial deviations from the straight line expected when distributions are similar, but we cannot assess how significant these deviations are. We turn to significance tests of the differences next.

We calculate the relative values separately for each type of laboratory, and we use the Kolmogorov–Smirnov equality of distributions test to see whether the distribution of laboratory procedures is different for physician office laboratories and independent laboratories (Table 4-5). In Table 4-5, we see that for four classes of procedures (Chemistry and Toxicology, Immunology, Microbiology, and Urinalysis), the submitted charges distributions are significantly different by laboratory ownership type. For all other classes of procedures, the relative values are not significantly different across laboratory ownership type, and for these we conclude that physician and independent laboratories have similar distributions. Because our analysis is exploratory, we do not attempt to explain this finding or draw any implications from it.

Table 4-5. Summary of Differences in Charge-based Relative Values Between Physician Office and Independent Laboratories

Class	Number of Procedures	Kolmogorov–Smirnov Test (p-value)
Chemistry and Toxicology	397	0.000
Cytogenetic Studies	23	0.548
Cytopathology	12	0.769
Drug Testing	3	0.996
Hematology	93	0.846
Immunology	149	0.000
Microbiology	133	0.001
Panels	7	0.432
Therapeutic Drug Assays	30	0.724
Transfusion Medicine	10	0.294
Urinalysis	10	0.029
Other	19	0.708

Finally, we identify those procedures with large differences in charge-based relative values between physician office laboratories and independent laboratories (Table B-4). For example, procedure 82000 (assay of blood acetaldehyde) under Chemistry and Toxicology has a difference of 694 percent. The associated relative values for physician office and independent laboratories are 5.84 and 0.74, respectively. This means that the relative value for the physician-based laboratory subset exceeds the independent laboratory-based subset by about 694 percent for this procedure. Differences in these observed values across laboratory types could reflect different methods used for testing and/or differences in submitted charges, among other things. It is beyond the scope of this analysis to explain these observed differences.

4.2.4 Regional Analysis

Policy makers may be interested in knowing whether absolute and/or relative laboratory costs differ between regions. True differences might warrant geographic adjustments to any new fee schedule. As argued above, the rationale for using charges to determine relative values is that charges may be systematically related to costs. While a laboratory may not have a definite idea about what the costs of performing a particular procedure are, they will likely have a good understanding of the relative cost of procedures. If costs are marked up to charges by a constant proportion across all procedures, then relative charges will do a good job reflecting relative costs. Then, a significant difference in the relative charges across two regions is suggestive of significant differences in relative costs across the two regions.

The data used in this analysis are aggregated by HCPCS code and by region using national adjusted volume as a weight. In computing the charge-based relative values, we used the class-based numeraire calculated separately for each region. Note that we used national adjusted volumes as weights when we calculated the region-specific numeraire because we did not want regional variations in volume to confound the analysis of relative prices. We calculated relative values by region for the 10 different geographic regions of the country identified in the claims data (Atlanta, Boston, Chicago, Dallas, Denver, Kansas City, New York, Philadelphia, San Francisco, and Seattle).

First, we present scatterplots comparing each region’s relative values to the national relative values. We cannot tell using the naked eye what constitutes significant deviation from the straight line—seeming outliers appear in almost every plot (Appendix C, Figures C-26 through C-35). However, the data appear to be more dispersed in Denver, New York, and Seattle, with more extreme outliers than in the other regions.

Next, we apply the Kolmogorov–Smirnov test to determine, on a pairwise basis, whether the distribution of charge-based relative values is significantly different in each region from the nation as a whole. Results from the Kolmogorov–Smirnov tests are presented in Table 4-6. We see that Chicago, Dallas, Denver, New York, and Seattle have schedules that are significantly different (at the 5 percent significance level) from the nation as a whole. We cannot tell from the analysis in Table 4-6 whether charge-based relative values are significantly higher or lower on average than the nation as a whole in these areas; all we know is that the underlying distribution of relative prices is significantly different in these regions.

Table 4-6. Summary of Regional Variation in Charge-based Relative Values

Region	Kolmogorov–Smirnov Test (p-value)
Atlanta	0.433
Boston	0.135
Chicago	0.035
Dallas	0.013
Denver	0.006
Kansas City	0.077
New York	0.015
Philadelphia	0.353
San Francisco	0.154
Seattle	0.007

5. COMPARISON OF VARIOUS PAYMENT SYSTEMS

5.1 Current Payments vs. a National Expenditure-Neutral Charge-Based Relative Value Fee System

In this section, we compare average fees based on the current fee system of allowed charges (which is the minimum of the NLA and local fee schedule amount) with fees based on a hypothetical expenditure-neutral system incorporating the charge-based relative values. To do this, we first calculate the average price per procedure under current payments as total expenditures per HCPCS code divided by procedure volume. This average is our estimate of the current allowed price (P_i^a) for each HCPCS code:

$$P_i^a = \frac{\text{Total Current Expenditures for Procedure i}}{\text{Total Volume for Procedure i}} \qquad (5.1)$$

Next, we establish a conversion factor that will make movement from the current payment system to the charge-based relative value system expenditure-neutral. The conversion factor (CF) is defined as the ratio of total current expenditures for all HCPCS codes divided by the sum of quantity-weighted (Q_i) charge-based relative values (RV_i) for each procedure:

$$CF = \frac{\text{Total Current Expenditures}}{\sum_i (RV_i * Q_i)}$$

for I = 1, 2,…n different test procedures. (5.2)

Thus, the new payment system is revenue-neutral because

$$\sum_i (RV_i\, Q_i) * CF = \text{Total Current Expenditures.}$$

With the conversion factor in hand, we then calculate the "converted" charge-based fee for each procedure (P_i^c) as follows:

$$P_i^c = CF * RV_i \qquad (5.3)$$

We then compare this expenditure-neutral fee (P_i^c) with the average fee in the current system (P_i^a). This allows for assessment of how fees for particular HCPCS codes would likely be affected by movement from the current mixed-rate system to an expenditure-neutral system with charge-based relative values.

We calculate two types of expenditure-neutral prices using Eqs. (5.1) through (5.3). The first is calculated as described above where the relative values are based on submitted charges with a universal numeraire based on all procedures. This leads to expenditure neutrality in the aggregate. The second type of price is calculated based on submitted charges with class-specific numeraires. In addition to being expenditure neutral in the aggregate, this set is also expenditure-neutral for each class of procedures. Table 5-1 shows the effect of class-based neutrality versus overall expenditure neutrality.

Column 1 shows current expenditures by class. Column 2 shows expenditures by class using charge-based relative values with a universal numeraire. For this column, there is overall expenditure neutrality, but there is not expenditure neutrality within classes—

that is, total expenditures increase (relative to what they actually are currently) for some classes and decrease for others. Column 3 shows the absolute difference in expenditures for each class, while Column 4 expresses the difference in percentage terms.

Using the universal numeraire and total expenditure neutrality results in a redistribution of expenditures between classes based on the relative values. This sort of redistribution was suggested in the sensitivity analysis in Section 4, where we looked at proportions of

Table 5-1. Summary of Expenditures, by Class, with Expenditure-Neutral Prices

	Current Expenditures ($1,000s)	**Expenditures with Charge-based Relative Values, Universal Numeraire ($1,000s)**	**Difference Between Columns (2) and (1)**	**Percentage Difference in Expenditures, By Class, Comparing Columns (2) and (1)**
Class: (column)	***(1)***	***(2)***	***(3)***	***(4)***
Chemistry and Toxicology	859,407	843,979	–15,428	–1.83
Cytogenetic Studies	4,621	2,791	–1,830	–65.59
Cytopathology	20,737	16,488	–4,249	–25.77
Drug Testing	3,514	2,444	–1,069	–43.74
Hematology	345,120	334,751	–10,369	–3.10
Immunology	87,912	98,092	10,180	10.38
Microbiology	130,484	131,962	1,478	1.12
Panels	236,099	235,326	–773	–0.33
Therapeutic Drug Assays	62,107	72,347	10,240	14.15
Transfusion Medicine	435	604	169	27.94
Urinalysis	64,631	75,945	11,313	14.90
Other	1,520	1,859	339	18.24
Total	1,816,587	1,816,587	0	0.00

relative values that would increase or decrease under a different valuation schedule. However, we could not be certain exactly how the redistribution would flow without looking explicitly at expenditures, as we do here. Column 4 displays the percentage difference between expenditures, by class, under the two different methods of accomplishing expenditure neutrality (total and class-based). The classes that would experience a decrease in payments using total expenditure neutrality are Chemistry and Toxicology, Cytogenetic Studies, Cytopathology, Drug Testing, Hematology, and Panels. However, Chemistry and Toxicology, Hematology, and Microbiology—the three largest classes—would actually see relatively little difference in actual reimbursements across methodologies (Column 3, Table 5-1). The classes that would experience a decrease in payments using class-based expenditure neutrality are Immunology, Microbiology, Therapeutic Drug Assays, Transfusion Medicine, Urinalysis, and Other. The largest percentage changes would occur for the smallest expenditure classes. The end result would be relatively small transfers between classes, as none of the amounts in Column 3 are very large. However, the impact on laboratories that specialize in cytogenetics, cytopathology, or drug testing could be large in terms of percentage revenue lost. Results for class-specific neutrality and total expenditure neutrality are not shown in Table 5-1, because there would be no redistribution of expenditures between classes with these assumptions.

5.2 Expenditure-Neutral Charge-Based vs. Expenditure-Neutral NLA-Based Relative Value Fee Systems

For this analysis, we obtain expenditure-neutral NLA-based relative values using computations similar to those described above for the charge-based expenditure-neutral prices (Section 5.1). To do this, we substitute the NLA-based relative values for the charge-based relative values (RV_i) in Eqs. (5.2) and (5.3). We then derive a second expenditure-neutral fee based on the NLA relative values, which we label P_i^{NLA}. As in the above analysis, we do this separately for the total neutrality and for the class-based neutrality. In Table 5-2, we compare these two expenditure-neutral price schedules, using both methods for computing the numeraire.

Table 5-2. Summary of Two Different Expenditure-Neutral Price Schedules, Using Two Different Methods for Computing the Numeraire for the Respective Relative Values

	Proportion of Procedures with:			
	Universal Numeraire		**Class-based Numeraire**	
	Charge-based (P_i^c) > NLA (P_i^{NLA})	**Charge-based (P_i^c) < NLA (P_i^{NLA})**	**Charge-based (P_i^c) > NLA (P_i^{NLA})**	**Charge-based (P_i^c) < NLA (P_i^{NLA})**
All Procedures	57%	42%	56%	43%
Class:				
Chemistry and Toxicology	55%	44%	56%	43%
Cytogenetic Studies	21%	79%	50%	42%
Cytopathology	8%	92%	25%	75%
Drug Testing	33%	67%	67%	33%
Hematology	76%	24%	78%	21%
Immunology	63%	34%	51%	47%
Microbiology	49%	51%	48%	51%
Panels[a]	43%	43%	57%	43%
Therapeutic Drug Assays	87%	13%	70%	30%
Transfusion Medicine	60%	40%	20%	80%
Urinalysis	50%	50%	30%	70%
Other	61%	39%	61%	39%

[a]Does not sum to 1 because for one procedure $P_i^c = P_i^{NLA}$.

By construction, the results for prices in Table 5-2 are identical to the results for relative values shown in Tables 4-1 and 4-2. Table 5-2 shows that average prices under the charge-based expenditure neutral fee system (P_i^c) would exceed average prices based on an expenditure-neutral NLA-based system (P_i^{NLA}) for more than 50 percent of procedures, using either the universal or the class-based numeraire.

5.3 Three-Way Comparison

In Table B-5, we list by HCPCS code five fee schedules: P_i^c based on an expenditure-neutral charge-based system using total neutrality, P_i^c based on an expenditure-neutral charge-based system using class-based neutrality, P_i^{NLA} based on an expenditure-neutral NLA-based system where the payment is based on the lower of the state fee schedule and the NLA, and P_i^a based on current allowed charges (this allows for the possibility that submitted charges are below the state fee schedule or the NLA). This table shows what the average price would be for each procedure code under the various expenditure-neutral pricing schemes. The published NLA price is also shown for comparison.

6. SUMMARY, DISCUSSION, AND AREAS FOR FUTURE WORK

6.1 Summary and Discussion

In our analysis, we compared submitted charge-based relative values to NLA-based relative values. In this section, we summarize the key results and discuss how the results can inform the possible development of an RVS-based fee schedule. We also identify areas for future work, based on our results.

Calculating relative values for nearly 1,000 clinical laboratory procedures necessarily produces an abundance of relative values for individual procedures, and amidst this abundance it is easy to lose track of the big picture. To focus our summary, we concentrate on the following general questions:

- Would charge-based relative values be significantly different from the relative values implied by the current NLAs?
- Does it matter if you use a universal numeraire instead of a class-specific numeraire to compute the relative values?
- Would an expenditure-neutral pricing system based on charge-based relative values change the distribution of reimbursements, relative to the current NLA-based system?
- If charge-based relative values were adopted, would utilization of individual procedures change?
- Did our analysis identify any other problems associated with charge-based relative values?

We answer each of these questions below.

Would charge-based relative values be significantly different from the relative values implied by the current NLAs?

Yes. Based on our statistical tests, the distribution of charge-based relative values is significantly different from the distribution implied by the NLAs. Moreover, looking at individual procedures, there are a number of large “outliers.” Outliers are defined as procedures with percentage differences (between the charge-based and NLA-based relative values) greater than one standard deviation from the mean differences. The standard deviation is large—63 percentage points around a mean difference of 19 percent—meaning that the outlier procedures would have prices at least 44 percent lower or 88 percent higher if charge-based relative values were used to set prices instead of the NLAs. These are substantial price differences.

Does it matter if you use a universal numeraire instead of a class-specific numeraire to compute the relative values?

Probably not. In principal, computing relative values with a universal numeraire for all classes of procedures allows for redistribution of expenditures across classes, even if prices are set to be expenditure-neutral in aggregate. In contrast, computing relative values with a class-specific numeraire leads to redistribution within but not across classes, assuming overall expenditure neutrality. The ability to redistribute expenditures across classes may be viewed as either an advantage or a disadvantage. It may be advantageous if overall prices are viewed as too high or too low in one class relative to another. However, various stakeholders may disagree about these judgments, and there does not appear to be any obvious empirical basis on which to make them in a systematic fashion.

In practice, our results suggest that the choice of numeraire does not have a large effect on the relative values we compute. For most procedures, the relative values computed with universal and class-specific numeraires are very close; they are almost always closer to each other than either of them is to the relative value associated with NLAs. The list of procedures that are outliers, relative to NLA-based relative values, is very similar for the universal and class-specific numeraires. The only large changes in outliers occur for Cytogenetic Studies, a relatively small class with many high-NLA tests.

Finally, the calculation of expenditures suggests that there will be relatively little redistribution between classes if a universal numeraire is adopted.

Would an expenditure-neutral pricing system based on charge-based relative values change the distribution of reimbursements, relative to the current NLA-based system?

This question can be answered at a variety of levels. At more aggregate levels, payments under an expenditure-neutral pricing system based on charge-based relative values would lead to relatively little redistribution across procedure classes, between independent and physician office laboratories, and across regions, as measured relative to the current NLA-based system. As noted in answer to the previous question, there will be relatively little redistribution between classes even if a universal numeraire is adopted. Similarly, Tables 6-1 and 6-2 show that the distribution of payments across laboratory types and by region, respectively, will be similar, whether prices are set on the basis of charge-based or NLA-based relative values.

However, when one moves to the individual procedure level, changing from NLA-based reimbursement to charge-based reimbursement would lead to large changes in reimbursement for individual procedures. These changes are directly related to the large differences between the charge-based and NLA-based relative values. Table B-6 shows total reimbursement for each procedure under charge-based and NLA-based relative values, assuming total expenditure neutrality. From this table, it is clear that there are large differences in total reimbursement for many procedures. Table 6-3 summarizes these differences by showing the percentage difference between charge-based and NLA-based reimbursement. These percentage differences are based purely on the difference in prices under the two systems. Looking at the center of the percent change distribution, about 18 percent of all procedures would see small percent changes in reimbursements (where the charge-based expenditure is within 10 percent of the NLA-based expenditure). These procedures account for about 42 percent of volume and 52 percent of current allowed charges. At the extremes of the percent change distribution, charge-based reimbursement will be more than 33 percent higher or lower than NLA-based reimbursement for about 43 percent of procedures. These represent very large percentage changes in reimbursement. However, these 43 percent of procedures account for less than 12 percent of current volume and allowed charges.

Table 6-1. Effect of Submitted Charge-based Prices on Revenue by Laboratory Type ($1,000s)

	Total Expenditure Neutrality	**Class-Based Expenditure Neutrality**	**Current Reimbursements**
Physician Office Laboratories	590,861	589,942	599,641
Independent Laboratories	1,212,634	1,213,685	1,204,027
Other Laboratories	13,092	12,960	12,918
Total	1,816,587	1,816,587	1,816,587

Table 6-2. Effect of Submitted Charge-based Prices on Revenue by Region ($1,000s)

	Total Expenditure Neutrality	**Class-Based Expenditure Neutrality**	**Current Reimbursements**
Atlanta	445,104	445,461	446,012
Boston	82,942	83,316	83,370
Chicago	244,942	244,597	242,902
Dallas	177,100	176,620	176,416
Denver	35,304	35,355	34,586
Kansas City	81,016	81,099	80,799
New York	293,810	294,060	294,250
Philadelphia	159,748	159,633	160,495
San Francisco	252,093	251,840	253,915
Seattle	44,526	44,606	43,843
Total	1,816,587	1,816,587	1,816,587

Table 6-4 provides a slightly different summary of the difference in reimbursement under charge-based and NLA-based relative values. The table shows the distribution of the difference in total reimbursement dollars under the two systems. The difference in total reimbursement dollars depends upon both the difference in price and the total volume for each procedure. Procedures can have small differences in expenditures if they have small differences in price or low volumes or both. The table shows that

33 procedures would have reimbursement of at least $1 million less under a charge-based system. While these 33 procedures are less

Table 6-3. Percentage Change in Total Reimbursements, in Moving from NLA-based to Charge-based Prices, by Numbers of Procedures and Percent of Procedures, Volume, and Allowed Charges

Charge-based Expenditures are	Number of Procedures	Percentage of Procedures	Percentage of Volume	Percentage of Allowed Charges
More than 67% lower	13	1.39	0.002	0.005
33% to 67% lower	136	14.53	4.88	4.27
10% to 33% lower	170	18.16	22.05	16.86
0% to 10% lower	91	9.72	32.49	40.68
0% to 10% higher	81	8.65	9.38	11.13
10% to 33% higher	180	19.23	24.56	22.64
33% to 67% higher	128	13.68	5.56	3.49
67% to 100% higher	51	5.45	0.72	0.59
More than 100% higher	86	9.19	0.37	0.36
Total	936	100	100	100

Table 6-4. Distribution of the Difference in Total Reimbursement Dollars, in Moving from NLA-based to Charge-based Prices, by Numbers of Procedures and Percent of Procedures, Volume, and Allowed Charges

Charge-based Expenditures are	Number of Procedures	Percentage of Procedures	Percentage of Volume	Percentage of Allowed Charges
More than $1,000,000 less	33	3.53	43.07	45.42
$100,000 to $1,000,000 less	65	6.94	13.06	13.32
$10,000 to $100,000 less	107	11.43	3.09	2.68
$1,000 to $10,000 less	97	10.36	0.18	0.37
$0 to $1,000 less	108	11.54	0.03	0.02
$0 to $1,000 more	114	12.18	0.02	0.03
$1,000 to $10,000 more	141	15.06	0.20	0.21
$10,000 to $100,000 more	157	16.77	1.75	2.12
$100,000 to $1,000,000 more	88	9.4	12.29	14.43
More than $1,000,000 more	26	2.78	26.31	21.41
Total	936	100	100	100

than 4 percent of all procedures, they account for more than 40 percent of total current volume and allowed charges, so the large change in reimbursement dollars is mainly attributable to the volume (rather than to the price changes per se). At the other extreme, 26 procedures (less than 3 percent of all procedures) would each receive reimbursement of at least $1 million more under a charge-based reimbursement system. Again, this large change in dollars is mostly attributable to volume—these procedures account for more than 26 percent of volume and more than 20 percent of current allowed charges. In the middle of the change in reimbursement dollars distribution lie the vast majority of procedures, accounting for only a small percentage of volume and charges. For these procedures in the middle, there is little change in reimbursement dollars.

Table 6-5 shows the procedures that have differences in reimbursement of at least $1 million. While these large changes could be disruptive if they were to occur all at once, the changes themselves could be beneficial in the long run. If the changes succeed in moving reimbursement prices closer to costs, this would benefit the industry. This scenario is not unlikely, because the current prices are based on such outdated information. To the extent that submitted charges better reflect current costs than prices that were in effect in the early 1980s, a shift toward charge-based reimbursement could help bring Medicare pricing closer to true laboratory costs. To ameliorate costs of adjustment, price changes could be transitioned gradually rather than enacted in one fell swoop.

Ideally, to test whether the charge-based prices better reflect costs than the current fee schedule, it would be necessary to have accurate cost information. Of course, if one had such information, it could easily be used to set administrative prices directly.

If charge-based relative values were adopted, would utilization of individual tests change?

Our calculation of expenditure-neutral prices assumes that the volume of utilization of individual tests will not change if we move from NLA-based reimbursement to reimbursement on the basis of

Table 6-5. Procedures with Reimbursement Changes in Excess of ±1 Million, Moving from NLA-based to Charge-based Prices

HCPCS Code	Description	NLA-based Reimbursements ($millions)	Charge-based Reimbursements ($millions)	Change in Reimbursements ($millions)	Percentage Change in Reimbursement
85610	Prothrombin time	73.5	92.4	18.9	0.26
83036	Glycated hemoglobin test	72.0	83.9	11.8	0.17
84439	Assay of free thyroxine	14.4	21.9	7.5	0.52
83970	Assay of parathormone	41.9	48.2	6.3	0.15
87340[a]	Hepatitis B surface ag	16.0	19.6	3.6	0.23
87086	Urine culture/colony count	29.5	33.0	3.5	0.12
82131	Amino acids	3.9	7.3	3.4	0.87
82108	Assay of aluminum	11.9	15.2	3.3	0.28
82043	Microalbumin quantitative	2.8	5.7	2.9	1.04
81001	Urinalysis auto w/scope	13.5	16.2	2.7	0.20
85652	RBC sed rate	2.7	5.4	2.7	1.00
82728	Assay of ferritin	23.9	26.4	2.5	0.10
80076[a]	Hepatic function panel	21.1	23.6	2.5	0.12
82270	Test for blood	6.9	9.1	2.3	0.32
81003	Urinalysis auto	4.5	6.6	2.1	0.47
85651[a]	RBC sed rate	9.7	11.7	2.0	0.21
G0107	CA screen; fecal blood test	5.5	7.3	1.7	0.33
84134	Assay of prealbumin	3.5	5.1	1.6	0.46
86140	C-reactive protein	3.7	5.3	1.6	0.43
85045	Reticulocyte count	3.1	4.6	1.5	0.48
87070	Culture specimen	6.9	8.1	1.1	0.17
80197	Assay of tacrolimus	1.1	2.2	1.1	1.00
80158	Assay of cyclosporine	2.2	3.3	1.0	0.50
84165	Assay of serum proteins	4.7	5.7	1.0	0.21
85022[a]	Automated hemogram	4.4	5.4	1.0	0.23
82784	Assay of gammaglobulin igm	5.4	6.4	1.0	0.19
80053[a]	Comprehensive metabolic panel	103.6	102.5	–1.1	–0.01
84478[a]	Assay of triglycerides	3.1	2.0	–1.2	–0.35
84295[a]	Assay of serum sodium	3.1	1.8	–1.3	–0.42

[a]Panel or procedure contained in panel or ATP.

(continued)

Table 6-5. Procedures with Reimbursement Changes in Excess of ±1 Million, Moving from NLA-based to Charge-based Prices (continued)

HCPCS Code	Description	NLA-based Reimbursements ($millions)	Charge-based Reimbursements ($millions)	Change in Reimbursements ($millions)	Percentage Change in Reimbursement
83718[a]	Assay of lipoprotein	6.2	4.9	–1.3	–0.21
84155[a]	Assay of protein	3.4	2.1	–1.4	–0.38
84153	Assay of PSA total	77.8	76.4	–1.4	–0.02
85023	Automated hemogram	15.8	14.3	–1.4	–0.09
82248[a]	Bilirubin	2.8	1.3	–1.4	–0.54
82310[a]	Assay of calcium	5.5	4.0	–1.4	–0.27
84550[a]	Assay of blood/uric acid	6.1	4.6	–1.5	–0.25
82977[a]	Assay of GGT	3.8	2.1	–1.7	–0.45
80054[a]	Comprehensive metabolic panel	25.3	23.5	–1.7	–0.07
84520[a]	Assay of urea nitrogen	8.7	6.9	–1.8	–0.21
84132[a]	Assay of serum potassium	9.5	7.7	–1.8	–0.19
87163	Special microbiology culture	5.1	3.2	–1.9	–0.37
83615[a]	Lactate (LD) (LDH) enzyme	5.3	3.3	–2.0	–0.38
82247[a]	Bilirubin	3.1	1.1	–2.0	–0.65
82985	Glycated protein	5.8	3.8	–2.1	–0.34
84075[a]	Assay alkaline phosphatase	4.0	1.7	–2.3	–0.58
82040[a]	Assay of serum albumin	4.0	1.6	–2.4	–0.60
86235	Nuclear antigen antibody	6.3	3.8	–2.4	–0.40
82947[a]	Assay of glucose	18.3	15.8	–2.5	–0.14
80162	Assay of digoxin	24.2	21.6	–2.6	–0.11
87186	Antibiotic sensitivity	13.8	11.0	–2.8	–0.20
84460[a]	Alanine amino (ALT) (SGPT)	9.6	6.6	–3.0	–0.31
87088	Urine bacteria culture	11.6	8.5	–3.1	–0.27
81000	Urinalysis nonauto w/scope	33.4	29.8	–3.6	–0.11
84450[a]	Transferase (AST) (SGOT)	9.5	5.9	–3.6	–0.38
82565[a]	Assay of creatinine	11.8	8.0	–3.8	–0.32

83721	Assay of blood lipoprotein	15.9	10.9	–5.0	–0.31

[a]Panel or procedure contained in panel or ATP.

(continued)

Table 6-5. Procedures with Reimbursement Changes in Excess of ±1 Million, Moving from NLA-based to Charge-based Prices (continued)

HCPCS Code	Description	NLA-based Reimbursements ($millions)	Charge-based Reimbursements ($millions)	Change in Reimbursements ($millions)	Percentage Change in Reimbursement
84443[a]	Assay thyroid stim hormone	165.4	158.9	–6.4	–0.04
85025[a]	Automated hemogram	121.7	111.8	–9.9	–0.08
85024	Automated hemogram	81.6	63.4	–18.2	–0.22

[a]Panel or procedure contained in panel or ATP.

charge-based relative values. This assumption may not hold. Volume is likely to change more in response to large price changes, and as noted previously, the magnitude of price changes implied by the charge-based relative values is often large. Patient and physician demand for Medicare laboratory tests may not be directly affected by price changes, because neither party pays for the tests. However, laboratories may change the way they market tests in response to large price changes and the resulting changes in profit margins for these tests.[10] For example, laboratories may market tests with increasing prices more aggressively, while cutting marketing for tests with lower prices. To the extent that the new prices better reflect underlying costs, the changes in marketing could increase market efficiency. However, it is difficult to predict the impact of changes in volume on aggregate laboratory allowed charges, because volume may rise for some tests and fall for others.

Did our analysis identify any other problems associated with charge-based relative values?

We found that calculating charge-based relative values may not help set reimbursement rates for automated test panels (ATPs).

[10]Lawsuits filed by the Department of Justice in the 1990s suggest that laboratory marketing could have significant effects on test volume. The lawsuits alleged that laboratories billed Medicare for unnecessary tests and persuaded physicians to order tests that the physicians thought were “free.” Such marketing efforts may be less likely in the current environment, because in settling the lawsuits, laboratories agreed to avoid such behavior in the future. In addition, there are now stronger requirements for showing medical necessity.

In the current fee schedule, selected automated chemistry tests are bundled for payment purposes, with reimbursement based purely on the number of tests performed and payments rising less than proportionately with the number of tests. The rationale for this approach is that the marginal cost of performing an additional automated test is less than the average cost of these tests. Our method for calculating charge-based relative values implicitly assumes that the cost and charges associated with a test are independent of other tests that may be provided at the same time as that test. In principal, the method could accommodate costs for bundles of tests, if submitted charges also reflect bundling (e.g., the charge for a bundle of tests A and B is less than the sum of the charge for A provided alone and the sum of the charge B provided alone). However, we conjecture that laboratories will submit the same charge for a procedure whether it is submitted alone or in a bundle. If this is the case, submitted charges are unlikely to provide useful information on the true cost relationship for ATPs.

Table 6-6 suggests the importance of setting prices accurately for panel procedures and procedures that are processed as ATPs. The table shows that panel-related tests are very common, accounting for about 40 percent of Medicare allowed charges. These tests account for nearly half of the tests that would have allowed charges change by over $1 million if charge-based relative values were used to set prices. Note that all but four of these tests would see reductions in allowed charges (see Table 6-5).

Table 6-6. Current Allowed Charges for Panel and Nonpanel Test Procedures

Procedure Type	Total Allowed Charges ($1000s)	Percentage of All Test Procedures
Nonpanel	1,082,832	59.60
Panel, component of an approved panel or ATP	733,755	40.39
Total	1,816,587	1.0000

6.2 Areas for Future Work

As is typical with studies of this nature, our results indicate a number of areas for future work if CMS decides to proceed with incorporating charge data into an RVS. Some of these areas are related to the limitations of our study, while other areas are related to policy development. We briefly discuss these areas below.

6.2.1 Validation

One of the limitations of the study is that there are no good data on relative costs to compare to the charge-based relative values. Therefore, there is no way to conclude that the charge-based relative values provide a "good" measure of relative costs, or that the charge-based relative values do a better job of measuring relative costs than the NLA-based relative values.

It might be possible to validate the charge-based relative values against costs by performing a detailed micro-costing study of laboratory procedures. Past attempts at laboratory micro-costing have proven difficult and expensive, in part because it is difficult to allocate aggregate laboratory costs across the many procedures produced in the laboratory, and in part because costs may be affected by laboratory scale. It is also possible that a micro-costing study might not identify the costs of efficient production if current reimbursement rates distort production decisions. Of course, if it is possible to successfully perform a micro-costing study, the results of the study could be directly used to develop an RVS.

An alternative form of validation for the charge-based relative values could focus on collecting additional data on the processes used by laboratories to set prices. This approach would investigate whether laboratories use common markups over cost when setting price, as our analysis assumes. This validation study might include a survey of laboratories. The success of the study would depend on laboratories' willingness to candidly respond about a sensitive subject: pricing strategy.

Another type of validation study could rely on expert opinion to compare the charge-based relative values to the NLA-based relative values and judge whether the charge-based relative values better reflect relative costs. Experts in laboratory production would carefully examine the relative values, based on their experience.

6.2.2 Applicability to Hospital Outpatient Laboratories

The results in this report are based on submitted charge data from independent laboratories and physician office laboratories from the Physician/Supplier Procedure Summary Master file. Summary charge data were not available from outpatient hospital laboratories, which account for about 40 percent of Part B laboratory allowed charges (IOM, 2000). Although summary submitted charge data for individual laboratory procedures are not available in Medicare outpatient data, existing data could be analyzed to determine whether use of a revenue-neutral fee schedule would lead to a change in reimbursement for hospital outpatient laboratories.

6.2.3 Transition

Our results show that changing from NLA-based pricing to a pricing system based on charge-based relative values could lead to large changes in prices for individual procedures. If CMS were to pursue such a charge-based approach, it might be advisable to gradually transition from the NLA levels to the new charge-based prices. Further development work could investigate how best to make the transition.

6.2.4 Panel Tests

We found that calculating charge-based relative values may not help set reimbursement rates for ATPs and their components. Because the components of ATPs are commonly performed, any further development work on a charge-based RVS should consider how to determine relative values for ATP components.

6.2.5 New Tests

Our analysis focused on pricing for laboratory tests currently covered by Medicare Part B. A complete reimbursement system will also need to consider how to set reimbursement rates for new tests. Future development work in this area could consider whether charge data should be incorporated into the process for setting prices for new procedures. On the one hand, charges for new procedures may provide useful information on appropriate fees for new tests. The charge data from claims for these procedures could be used to set new fee schedule amounts without requiring the gap-filling or cross-walking that is currently used to set fees. On the other hand, laboratories may submit higher charges if they believe that submitted charges will be used to determine fees for new tests.

REFERENCES

Health Care Financing Administration (HCFA). March 1997. Program Memorandum Intermediaries/Carriers. HCFA Publication 60AB, Rev. AB-97-5. Subject: New Panels Approved by CPT—Action. <www.hcfa.gov/pubforms/transmit/ab%2097%2D5.htm>.

Health Care Financing Administration (HCFA). November 1999. Program Memorandum Intermediaries/Carriers. Transmittal AB-99-85. Subject: Clinical Diagnostic Laboratory Organ or Disease Panel Codes Claims Processing Procedures for April 2000. <www.hcfa.gov/pubforms/transmit/ab998560.htm>.

Institute of Medicine (IOM). 2000. *Medicare Laboratory Payment Policy: Now and in the Future*. Washington, DC: National Academy Press.